Assignment 1: NYC Real Estate Expansion: Analytics for Location and Specialization

By

Oluwaseyi (Olu) Olayeye

Boston University Metropolitan College

AD571: Business Analytics Foundations

05/22/2023

Table of Contents

[Executive Summary 3](#_Toc135744062)

[Problem Understanding and Definition 4](#_Toc135744063)

[Data Requirement Identification and Collection 5](#_Toc135744064)

[Data Cleaning and Preparation 5](#_Toc135744065)

[Exploratory Data Analysis 5](#_Toc135744066)

[Five – Number Summary 5](#_Toc135744067)

[Total Sales by Property Type – Boroughs, Neighborhoods 6](#_Toc135744068)

[Sales Distribution in 2021 – NYC, Queens, Richmond Hill 6](#_Toc135744069)

[Sales Transaction Expected Growth Rate 7](#_Toc135744070)

[Compound Annual Growth Rate (CAGR) 7](#_Toc135744071)

[Total Revenue 8](#_Toc135744072)

[Appendix-1 10](#_Toc135744073)

[Figure 1: 5 – Number Summary – Property Type 10](#_Toc135744074)

[Figure 2: 5 – Number Summary – Gross Square Feet & Sale Price 10](#_Toc135744075)

[Appendix-2 11](#_Toc135744076)

[Figure 3: Total Sales by Property Type – Boroughs, Neighborhoods 11](#_Toc135744077)

[Figure 4: Sales Distribution in 2021 – Queens-Richmond Hill 11](#_Toc135744078)

[Appendix-3 12](#_Toc135744079)

[Figure 5: Sales Distribution in 2021 – NYC-Richmond Hill 12](#_Toc135744080)

[12](#_Toc135744081)

[Figure 6: Five-Year Expected Growth Rate (2017 – 2021) – Queens-Richmond Hill 12](#_Toc135744082)

[Appendix-4 13](#_Toc135744083)

[Figure 7: Five-Year Expected Growth Rate (2017 – 2021) – NYC-Richmond Hill 13](#_Toc135744084)

[Figure 8: Compound Annual Growth Rate (CAGR) 13](#_Toc135744085)

[Appendix-5 14](#_Toc135744086)

[Figure 9: Total Revenue 14](#_Toc135744087)

[Figure 10: Heatmap Sale Price – Queens 14](#_Toc135744088)

# Executive Summary

Our Exploratory Data Analysis (EDA) of NYC real estate markets, particularly Richmond Hill in Queens, revealed key trends and insights. Residential sales lead the market, accounting for $145.3 billion in Queens, with Richmond Hill standing out at $4.56 billion. Sales distribution in 2021 showed a significant upsurge during the summer months, hinting at heightened real estate activities during this period. We also observed a positive correlation between property size and sale value, with larger, high-end properties greatly influencing the average prices.

The EDA also assessed the five-year sales transaction growth rate, Compound Annual Growth Rate (CAGR), and total revenue. Positive sales transaction growth trend was noticeable from 2017-2021 across NYC, Queens, and Richmond Hill. Queens and Richmond Hill demonstrated a CAGR of 4.96% and 4.76% respectively from 2003-2021, indicating steady growth over this timeframe. Richmond Hill registered a notable total revenue of $1.7 million in 2021, ranking second in Queens borough.

Our analysis suggests that Richmond Hill is a good new real estate office location. We must examine additional factors in more detail to make an informed decision. Specifically, we should look at metrics like property age, sale price per square foot, sale price per unit, and days on the market. These data points can give us deeper insights into the local real estate market. Additionally, we should consider operational costs, potential competition, and how our strategic objectives align with company goals. While we can use growth predictions and the CAGR to guide our analysis, we should be cautious since past performance doesn't always predict future results. Considering all these factors, we can determine if Richmond Hill is the right place to establish our new real estate office.

# Problem Understanding and Definition

The objective of this project is to provide data-driven insights that will guide the strategic decision-making of a real estate brokerage firm planning to open an office in New York City. The primary question to be answered is: Which neighborhood should the firm choose for their office location to maximize opportunities and minimize risks?

# Data Requirement Identification and Collection

For this analysis, we require the NYC Real Estate data to help us understand the real estate market dynamics, property characteristics, and sales trends in various neighborhoods of NYC.

We will connect to SQL Server from Power BI to extract the required tables for our analyses.

# Data Cleaning and Preparation

The initial stage of our analysis involves data cleaning and preparation in Power BI. This process will help maintain data quality and accuracy for subsequent analytical procedures. The steps include:

1. Filtering out invalid year entries in the Year\_Built field.
2. Removing properties with 0 Gross\_Square\_Feet, which can skew size calculations.
3. Excluding properties with Sale\_Price values below $10,000, which are likely errors.
4. Removing rows with blank Building\_Code\_ID, corresponding to Building Type.
5. Converting ID fields to text format to prevent unwanted numerical operations.

# Exploratory Data Analysis

## Five – Number Summary

Our EDA focused on comparing residential and commercial property transactions across various neighborhoods and boroughs, emphasizing Richmond Hill in Queens. The analysis revealed that residential transactions involving smaller properties were more common across all areas. However, commercial transactions, while less frequent, had potential for larger property sizes in particular regions, such as Jamaica in Queens. ([Figure 1](#_Figure_1:_5))

Richmond Hill typically features smaller properties in terms of Gross Square Feet (GSF), translating to lower sale prices compared to neighborhoods with larger properties. Nonetheless, differences are apparent across boroughs, with Manhattan featuring significantly larger properties and higher sale prices. ([Figure 2](#_Figure_2:_5))

A clear correlation between Gross Square Feet and Sale Price emerged, pointing to the fact that larger properties command higher prices. This finding is important, as areas with higher prices could present profitable opportunities due to their direct impact on revenue potential and indication of market activity.

## Total Sales by Property Type – Boroughs, Neighborhoods

Borough-wise, Manhattan led in total sales for commercial and residential properties, followed by Brooklyn, Queens, and the Bronx. However, in Queens, residential sales significantly exceeded commercial sales, suggesting a predominantly residential real estate market, a trend observed across other boroughs except for Manhattan.

Neighborhood-wise, Richmond Hill registered the highest total residential sales in Queens and showed significant commercial and mixed property sales, indicating a diverse real estate market. Although other neighborhoods like Forest Hills and South Jamaica also had high residential sales, their commercial and mixed property sales were lower. Jamaica stood out with high sales across all property types. ([Figure 3](#_Figure_3:_Total))

## Sales Distribution in 2021 – NYC, Queens, Richmond Hill

For NYC and Queens, sales transactions peaked in August and were lowest in December, revealing a seasonal pattern. ([Figure 4](#_Figure_4:_Sales))

However, Richmond Hill showed a more consistent sales distribution across the months. The neighborhood saw smaller peaks in June and August but maintained a steady transaction rate throughout the year. ([Figure 5](#_Figure_5:_Sales))

In conclusion, while NYC and Queens show a clear seasonality pattern in sales volume, Richmond Hill demonstrates less seasonality, maintaining consistent transactions throughout the year. Richmond Hill's steady sales and alignment with broader Queens and NYC market trends make it a potential candidate for a new real estate office. However, other variables, such as local market conditions, property prices, and competitive landscape, must be considered before making the final decision.

## **Sales Transaction Expected Growth Rate**

NYC and Queens witnessed significant growth in transaction numbers from 2017 to 2021, despite a downturn in 2020, likely due to the pandemic. Specifically, NYC transactions increased from 27,270 to 60,537, while Queens grew from 10,623 to 17,133. ([Figure 6](#_Figure_6:_Five-Year))

Contrastingly, Richmond Hill maintained stable transaction numbers ranging between 307 and 438 during this period, despite experiencing a downturn in 2020 and a slight recovery in 2021.

In conclusion, while NYC and Queens showed substantial growth in sales transactions, Richmond Hill demonstrated resilience and stability. Hence, Richmond Hill presents potential as a suitable neighborhood for a new real estate office, given its steady performance and alignment with the Queens and NYC market trends. However, considerations of local demand, supply conditions, property prices, and market competition remain crucial factors in making a final decision. ([Figure 7](#_Figure_7:_Five-Year))

## **Compound Annual Growth Rate (CAGR)**

Among all boroughs, Brooklyn exhibited the highest CAGR at 6.73%, while Manhattan showed a negative growth rate at -3.34%. The Queens borough displayed moderate growth at 4.96%. Within Queens, Jamaica had the highest neighborhood CAGR at 5.44%, while Forest Hills experienced the lowest growth at 0.82%.

Applying the CAGR to predict the expected number of transactions for 2022, Queens could have approximately 23,338 transactions, and Richmond Hill around 400.

Despite its low growth rate, Forest Hills emerged as a high-value neighborhood, indicated by the largest maximum gross square feet and a high maximum sale price among the examined Queens neighborhoods. ([Figure 8](#_Figure_8:_Compound))

In conclusion, Richmond Hill shows robust growth, aligning closely with the broader Queens borough growth trend, indicating it could be a suitable neighborhood for a new real estate office. However, these predictions are based on past trends and must be evaluated alongside factors like local demand, supply conditions, property prices, and market competition, as future market conditions may vary.

## **Total Revenue**

This analysis examined total revenue for 2021 with revenue calculated based on a 5% commission on residential sales and 12.5% market penetration.

Borough-wise, Manhattan leads with over **$268 million** in revenue, attributed to its high-value real estate market. Brooklyn follows with approximately $125 million, while Queens ranks third, generating around $90 million. These numbers indicate lower property prices in Queens despite a significant number of transactions. Staten Island and the Bronx have about **$25 million** in comparable revenues, considerably lower than the other boroughs.

Forest Hills generates the highest revenue among Queens' neighborhoods, over **$3 million**, consistent with earlier analyses indicating larger, high-end properties. Richmond Hill is the second-highest revenue earner, with around $1.7 million, reflecting its strong real estate market. In contrast, Jamaica Bay has the lowest revenue of about **$50k**, suggesting fewer transactions, lower-priced properties, or both. ([Figure 9](#_Figure_9:_Total))

Comparing these revenue figures with sales trends observed in previous analyses, higher sales volume does not necessarily translate into higher revenue, particularly in areas where property prices are lower. The revenue seems to align more closely with the property prices in the areas than with the number of transactions.

The standout point here is the significant disparity in revenue between boroughs and neighborhoods. It's interesting to see how the property prices and market activity in each borough or neighborhood directly affect the revenue generated. Such insights can be crucial for real estate companies when planning their strategies or deciding where to open new offices.

# Appendix-1

### Figure 1: 5 – Number Summary – Property Type

A picture containing text, screenshot, number, font

Description automatically generated

### Figure 2: 5 – Number Summary – Gross Square Feet & Sale Price

A picture containing text, screenshot, number, font

Description automatically generated

# Appendix-2

### Figure 3: Total Sales by Property Type – Boroughs, Neighborhoods

A picture containing text, screenshot, font, diagram

Description automatically generated

### Figure 4: Sales Distribution in 2021 – Queens-Richmond Hill

A picture containing text, screenshot, diagram, line

Description automatically generated

# Appendix-3

### Figure 5: Sales Distribution in 2021 – NYC-Richmond Hill

### A picture containing text, screenshot, rectangle, line Description automatically generated

### Figure 6: Five-Year Expected Growth Rate (2017 – 2021) – Queens-Richmond Hill

A picture containing text, line, diagram, plot

Description automatically generated

# Appendix-4

### Figure 7: Five-Year Expected Growth Rate (2017 – 2021) – NYC-Richmond Hill

A picture containing text, line, diagram, plot

Description automatically generated

### Figure 8: Compound Annual Growth Rate (CAGR)

A picture containing text, screenshot, font, number

Description automatically generated

# Appendix-5

### Figure 9: Total Revenue

A picture containing text, screenshot, font, number

Description automatically generated

### Figure 10: Heatmap Sale Price – Queens

A picture containing text, screenshot

Description automatically generated